Dr. G. Y. Pathrikar College of Computer Science and Information Technology,

Chhatrapati Sambhajinagar

Final Synopsis Entitled

"Extraction and Classification of Skin Lesion towards Melanoma Skin Cancer Using Machine Learning Techniques"

Submitted by

Ms. Apurva Sureshrao Solanke

Ph.D Student

Under the Guidance of

Dr. Prapti Devidas Deshmukh

Principal, Dr. G. Y. Pathrikar College of CS & IT, MGM University

Submitted To

Dr. G. Y. Pathrikar College of Computer Science and Information Technology,

MGM University, Chhatrapati Sambhajinagar, Maharashtra

"Extraction and Classification of Skin Lesion Towards Melanoma Skin Cancer Using Machine Learning Techniques"

1. Introduction

Skin plays an important role in regulating body temperature and in protecting us from the environment, including sunlight. Human skin cancer is potentially life-threatening form of cancer. Skin cancer is one of the most common types of cancer all over the world. It is easily treatable when it is detected in its beginning stage. There are different types of skin cancer, among all types of skin cancer, melanoma is one of the deadliest types of skin cancer in human that appear on the skin as pigmented moles or marks. Melanoma skin cancer is one of dangerous types of cancer because it's much more likely to spread to other part of the body if not diagnosed and treated early. Melanoma is one of the most malignant and dangerous types of skin cancer that causes a majority of deaths related to skin cancer. The no of cases of melanoma skin cancer has been increasing year by year. Due to this, it is necessary to examine and observe melanoma closely when it is still at the early stage. Early detection of melanoma is important in reducing the mortality rate of skin cancer. Melanoma is less common than basal cell and squamous cell skin cancer but it is more dangerous than other two types. Malignant melanoma, often known as melanoma, is a highly aggressive form of skin cancer. Traditionally, melanoma is detected through painful and time-consuming biopsies. Skin cancer is usually diagnosed by biopsy, but the process is slow and not always accurate. Recently machine learning becomes an efficient method in classifying skin lesions as melanoma or benign. Main features for this task include colour, texture and shape. To improve the accuracy and efficiency of melanoma detection, the integration of advanced computational techniques such as machine learning (ML) in medical imaging has gained significant attention. This research focuses on the extraction and classification of skin lesions using machine learning techniques to aid in the early detection of melanoma. By processing dermatological images, the aim is to develop a robust and automated system that can assist healthcare professionals in identifying melanoma skin cancer with greater accuracy. In research we have applied Preprocessing techniques, Segmentation techniques, Different feature extraction, classification techniques. In preprocessing we have used point processing, contrast stretching, image negative, Laplacian filter and some other filters also but we get best result in morphological closing and histogram equalization for noise removal and to enhance other features. While segmentation we have applied k mean clustering, Otsu thresholding, watershed segmentation, edge-based segmentation but we get the good result in Otsu thresholding for the purpose of separation of background from foreground. We used segmentation to get lesion area separately for further processing. After segmentation we have used feature extraction to extract exact feature for further classification. We used ABCD rule for the purpose of extraction of features. After that we have applied machine algorithms for classification of skin cancer whether it is benign of melanoma. We get the best accuracy in SVM classification algorithm.

2. Research Objectives

- To **extract meaningful features** from skin lesion images for melanoma detection.
- To develop an **automated classification system** that differentiates between melanoma and non-melanoma lesions.
- To design a system that can be used for **early detection** of melanoma in clinical settings.

3. Literature Review

Arora et al. developed computer-aided detection and diagnosis systems for classifying a lesion into cancer or non-cancer. They proposed feature extractor and quadratic support vector machine for skin lesions classification. Te PH2 dataset was used, and the model achieved accuracy of 85.7%, sensitivity of 100%, specificity of 60% and training time of 0.8507 S [1]. Senan et al. applied the ABCD (Asymmetry, Border, Colour and Diameter) rules for automatic skin cancer detection. Tey used PH2 standard dataset. Gaussian filter was applied to enhance the images. the contour method was applied for extracting the Region of Interest (Roi) from dermoscopy images. Morphology method was applied for increasing the quality of skin lesions. the ABCD rules were implemented for features extraction. Accuracy, specificity and sensitivity of the system were calculated. Accuracy was 84%, specificity was 89.50% and sensitivity was 60.50% [2]. (M. Krishna Monika) In this study author discovered that there is a significant increase in the number of skin cancer patients due to a variety of reasons. As a result, early detection is critical in both detection and treatment. Thus, this study discusses a strategy based on the MSVM classification, which employs two effective methods for feature extraction known as ABCD and MSVM [3]. A technique is being developed to assist patients and doctors in detecting or identifying skin cancer classifications, whether benign or malignant. MA. Ahmed Thaajwer and UA.Piumi Ishanka (2021) recommended using the ISIC dataset, which comprises of23,000 images, 600 of which are used for training and testing [4]. In this work proposed a project to examine the accuracy of skin cancer prediction as well as the classification of skin cancer as malignant or non-malignant melanoma. To accomplish this, certain pre-processing steps were performed, including hair removal, shadow reduction, glare removal, and segmentation. To categorise, SVM and Deep Neural networks will be employed [5]. This paper proposes an idea to classify the melanoma using shearlet transform coefficients and naïve Bayes classifier. The dataset is decomposed using shearlet transform with the predefined number of (50, 75 and 100) shearlet coefficients. Then to the naïve bayes classifier, the required coefficients are applied. The accuracy achieved at 3rd level of classification using 100 coefficients of shearlet transform [6]. In 2015, Amruta M. Gajbar and Prof. A. S. Deshpande, with a study entitled" Detection And analysis of skin cancer in skin lesions by using segmentation", where 5 features (contrast, correlation, energy, homogeneity, and standard deviation) were extracted and used as input to grayscale co occurrence matrix (GLCM) algorithm, and there were be able to classify output to Basal cell carcinoma and melanoma [7]. In 2015 V. Jeya Ramya, J. Navarajan, R. Prathipa, and L. Ashok Kumar conducted a study entitled " Detection of melanoma skin cancer Using Digital Camera Images' in which the same

algorithm was used in the previous study, but the output was classified into a benign mole and melanoma [8]. In 2015, with a study entitled "Feature extraction for skin cancer lesion detection", in which they made use of the ABCD rule developed by the American Cancer Society to identify melanoma. The values of the four characteristics were extracted: Asymmetry, Border, Colour, Diameter, and then TDV is calculated by the formula: TDV = A 1.3 + B · 0.1 + C 0.5 + D · 0.5. By the TDV value, the output is classified into a benign mole, suspected mole, or melanoma [9]. In this research, a total of 2158 images are taken of examples such as a typical Nevus, common nevus and melanoma skin cancers. As the dataset is large the resulting accuracy may also depend on the quality of capturing the photos. Using CNN (convolutional neural networks) and SVM (support vector machine) algorithms to identify the type of skin cancer the patient is affected with and achieved an accuracy of 92% and 95% respectively [10]. In this paper, hybrid feature extraction has been used to classify skin lesion either as benign or melanoma. Automatic detection of skin lesion with ABCD rule, GLCM and HOG for feature extraction and classification using machine learning techniques. GAC method was proposed for segmentation of the skin lesion. Segmentation result with 0.9 JA and 0.82 DI has been achieved. For colour, symmetry and diameter of skin lesion ABCD rule, texture of skin lesion GLCM and shape, edge of skin lesion HOG was proposed for feature extraction. Different machine learning techniques such as SVM, KNN and Naïve Bayes was proposed to address the classification. The classification result obtained is 97.8 % of Accuracy and 0.94 Area under Curve using SVM classifiers. And additionally, the Sensitivity obtained was 86.2 % and Specificity obtained was 85 % using KNN [11]. In this paper an effective detection of skin cancer cells is proposed. Four features are chosen that are trained tested by using various classification techniques like SVM, KNN, Decision tree and Boosted tree have been done. The result discusses that SVM is a better classifier than KNN, DT, BT. SVM has the highest accuracy of 93.70% [12]. In this research experience shows that only pre-processing, segmentation and feature extraction are not sufficient to detect a skin lesion. Classification of skin cancer lesions is done with PNN classifier to classify a skin lesion are cancerous or normal [13]. In this project, author used the dataset of 4000 images from well-known ISIC for checking out and training. The cancer detection approach is completed in precise levels like pre-processing, segmentation, feature extraction, and classification. Based on the all features like colour, texture, and shape features along with trained data, texture features are done by the Gabor filter. SVM will predict skin cancer whether it is malignant or benign. In order to detect and classify skin cancer for a single image, author used a Support Vector Machine (SVM) for the detection and classification of images [14]. In this study CNN model performs well in image classification tasks like classification of skin cancer. To improve the model's stability and convergence, we pre-processed the dataset by dividing it into a training and test set, oversampling the minority classes to solve class imbalance, and normalising the pixel values. Finally, author calculated the accuracy and loss of the model on both the training and validation sets to assess its performance. Findings demonstrated that model performed better than the standard random guessing approach, obtaining a test accuracy of almost 96% [15]. In this paper, an ABCD technique is proposed to detect the malignant melanoma at an early stage in order to reduce the medical cost of taking biopsy. an ABCD technique is proposed to detect the melanoma skin cancer. This study narrates the procedure and methodologies of image processing and soft computing techniques used to diagnose the melanoma with better accuracy. ABCD technique provides differentiation between the normal moles, benign and malignant

melanoma and also used to identify malignant melanoma at an early stage without taking biopsy [16]. The various methods for melanoma detection mentioned in this survey research are more effective at detecting the disease from images of skin lesions than clinical diagnosis. To get accurate results, the four stages of the melanoma diagnosis process are carried out using contemporary methods. The investigation taken into consideration a number of methods, including the rules for regulating GA, SVM, ANN, and ABCD. Early melanoma identification will be possible when these techniques are combined and applied to skin lesion images. In order help in the classification of skin lesions, the significance of a programmed characterization technique is investigated in this work [17]. The paper addressed melanoma cancer detection using machine learning including the steps of preprocessing, segmentation, feature extraction and classification. In this methodology original color picture is chosen from dataset. Chosen colored skin picture is changes over into gray shading picture[1]. Skin picture contains a few hairs which will corrupt the precision of results. Hair expulsion is needed and is done by using Gaussian filter. Affected part is required for picture investigation with the goal of segmentation is performed. Finally, author proved that using skin image we are able to classify the image successfully with the maximum accuracy, the implementation was mostly achieved using Machine Learning and Image processing techniques [18]. This paper aims to develop a simple method capable of detecting and classifying skin lesions using dermoscopy images based on ABCD rules. The proposed approach is implemented in the MATLAB environment and the experiment is based on PH2 database containing suspicious melanoma skin cancer. This method classifies images with a specificity of 92% and a sensitivity of 87% which reflects its reliability [19]. The main goal of this research is to eradicate human involvement in melanoma cancer diagnosis, making it less error prone and time consuming. To do some pre-processing steps were carried out which are Hair removal, Glare removal and segmentation. After pre-processing images is segmented using maximum entropy method. Greatest entropy thresholding is utilized to discover local of interest [20]. The proposed method, using the SVM was successful. The result shows accuracy of 96.9% of detecting melanoma disease with other disease. The results explain the performance evaluation of SVM [21]. The visual results are presented in the description and the accuracy of mathematical analysis is 94.88 percent. This system will give more accuracy and will generate results faster than the traditional method, making this application an efficient and dependable system for dermatological disease detection. Furthermore, this can also be used as a reliable real time teaching tool for medical students in the dermatology stream [22]. In this research work, The ABCD algorithm had been proposed and implemented in the skin lesion images from the PH2 dataset and used TDS as the classifier. In this work, we depend on a pre-processing step to hair removal. It based on morphological filter and thresholding. The ABCD algorithm was applied to the image with the objective of feature extraction to detect melanoma. These features include structural properties of shape and irregular distribution of colour. This method achieved an accuracy, specificity and sensitivity in the range of 93.2%, 92.59%, and 90.15% respectively [23]. In this paper, author presented an automated method to segment skin lesions regions using a semi-supervised learning technique. Author uses two major steps, namely pre-processing and segmentation. During pre-processing, artefacts such as air bubbles, hair and ink marks are filtered out. Lesion regions are segmented using Grab Cut and clustering algorithms. The Grab Cut algorithm segments the foreground image (lesion region) and k-means clustering

and flood-fill algorithms are applied to obtain the lesion region with improved boundaries. To ascertain the performance of the proposed algorithm we have tested our method with ISIC 2017 dataset and PH2 dataset for both pre-processed and non-pre-processed images [24]. This paper represents efficient method for skin cancer detection. In this system skin image is pre-processed then segmentation is performed using a segmentation technique which is followed by feature extraction. Classification is performed on extracted features. SVM classifier is applied for classification. It is observed that SVM linear function perform better than radial basis function and Bayes net classifier. SVM linear function gives highest accuracy for the classification of melanoma from benign lesion [26]. Melanoma identification necessitates a number of procedures, including pre-processing, segmentation, labelling, and feature extraction. This article focused on various approaches such as Hybrid Artificial Neural Networks, Genetic Algorithms, Neuro Fuzzy Systems, Hybrid Genetic Algorithm and Artificial Neural Networks, SVM, CNN, ABCD law, and unsupervised algorithm – K means algorithm, among others. However, of all the algorithms studied, the SVM algorithm approach is the best for cancer detection because it has the fewest drawbacks. We assume that the neural network technique is the strongest of all the available methods [27].

4. Dataset

A dataset for melanoma skin cancer is a structured collection of medical and dermoscopic images often accomplished by clinical metadata and diagnostic labels used for research, diagnosis and development of automated detection systems. A melanoma skin cancer dataset is an organised repository of dermoscopic or clinical skin images, primarily used for studying melanoma and other skin lesions, often annotated by dermatologists with ground truth labels such as benign or malignant. There are generally two types of datasets used for melanoma image analysis: local and global. Local datasets consist of medical images collected from specific hospitals, clinics or regional healthcare centres. Access to such datasets is usually restricted due to privacy concerns and strict regulations regarding patient confidentiality. Because of these privacy and ethical limitations, we were unable to obtain or use a local dataset for our study. As a result, we relied on a global dataset, which is publicly available and compiled from various international sources. Global datasets are commonly used in research because they include a wide variety of images collected from diverse populations around the world.

4.1 Global datasets

A global dataset is a large, publicly available collection of data shared internationally, often managed by research groups, universities or health organization. These datasets are open to access online with contributions from multiple countries and are designed for broad use and standard benchmarking. They usually contain thousands of well-labelled images with diverse demographics and clinical details. Global datasets are high-quality and reliable, making them ideal for training strong AI models. They also support comparisons between studies and serve as standard benchmarks in research. Several publicly available skin cancer datasets are used for research and

development in the field of skin cancer detection and diagnosis. These datasets include images of various skin lesions, both benign and malignant and are often used to train and evaluate machine learning models.

Recently, Kaggle launched an interesting competition to identify melanoma in images of skin lesions. Skin cancer is the most prevalent type of cancer. Melanoma, specifically, is responsible for 75% of skin cancer deaths.

In the context of melanoma skin cancer research, databases play a vital role in storing and managing clinical and imaging data used for diagnosis, segmentation, classification, and machine learning. Below are the types of databases commonly used, along with detailed explanations and examples:

1. **ISIC (International Skin Imaging Collaboration)**

The International Skin Imaging Collaboration (ISIC) is a global effort between researchers and industry to reduce skin cancer deaths using digital skin images. It promotes standards for skin imaging and brings together experts in dermatology and computer vision to improve diagnosis. ISIC supports early melanoma detection by sharing dermoscopic image datasets, hosting challenges, and providing open-source tools.

2. **HAM10000 (Human Against Machine with 10,000 training images)**

HAM10000 ("Human Against Machine with 10,000 images") is a large, curated dataset of dermoscopic images used for skin cancer research, especially melanoma detection. It's widely used for tasks like classifying melanoma vs. benign lesions, segmentation, data augmentation, transfer learning, and handling imbalanced data. The dataset is standardized, covers key pigmented skin conditions, and includes metadata like age, sex, and lesion location for deeper analysis.

3. **PH2 database-**

The PH2 dataset is a dermoscopic image database created for research and benchmarking in skin lesion segmentation and classification. It was collected at Hospital Pedro Hispano in Portugal and contains 200 images: 80 common nevi, 80 atypical nevi, and 40 melanomas. Each image includes medical annotations such as lesion segmentation, clinical and histological diagnoses, and dermoscopic features.

5. Methodologies

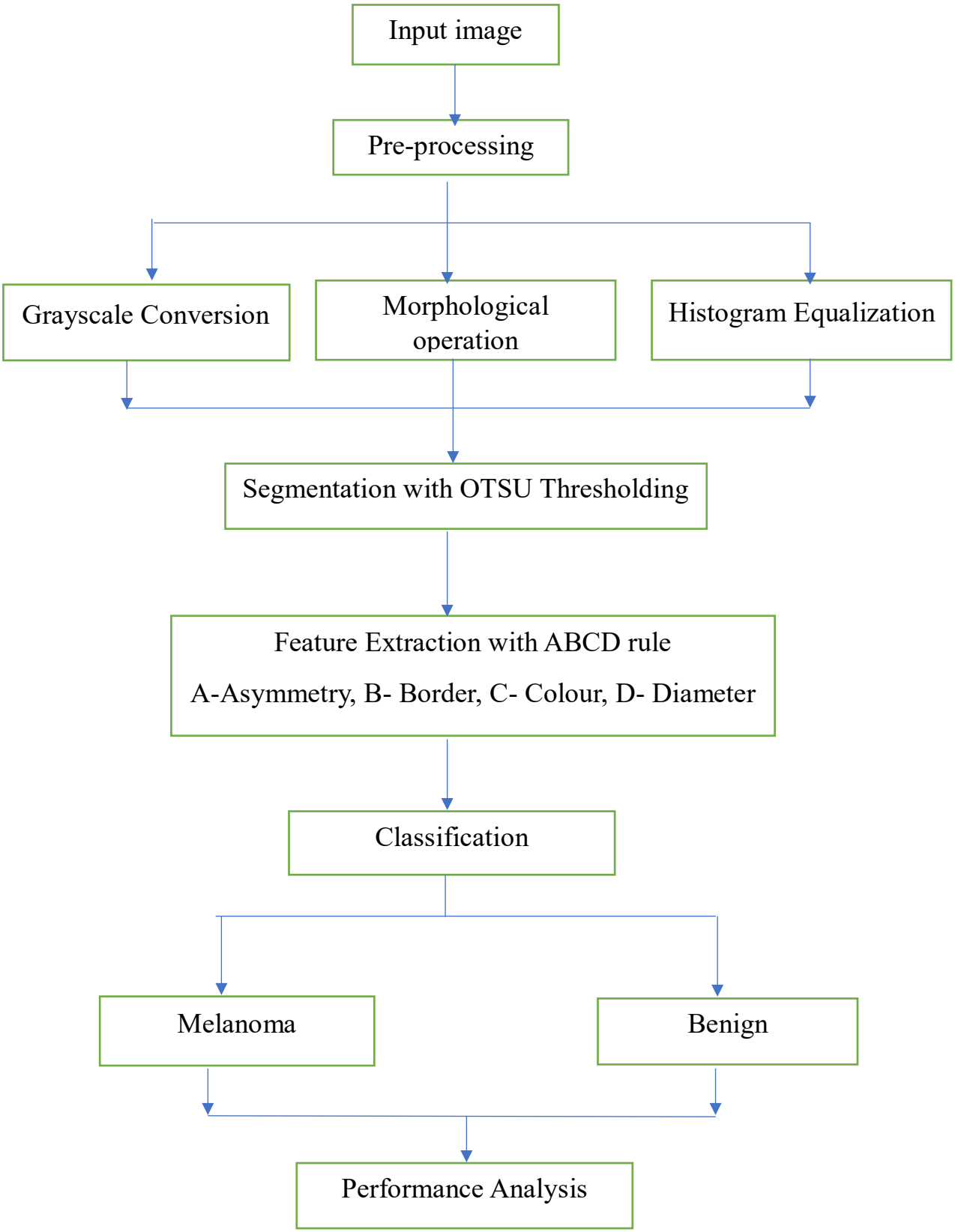


Figure.1 System Workflow diagram

5.1 Input image

In this research, image acquisition from local sources was not possible due to privacy and data protection regulations. As a result, we utilized publicly available global datasets for this research. The datasets used include the ISIC (International Skin Imaging Collaboration), PH2 datasets, which provide high-quality dermoscopic images essential for melanoma classification. These datasets offer diverse and well-annotated images, ensuring reliable training and evaluation of machine learning models.

5.2 Preprocessing

Preprocessing is essential to remove artifacts such as hair, ruler marks, and uneven lighting from dermoscopic images. In this study, we applied various techniques including contrast stretching, Laplacian filter, median filter, Gaussian filter, image negative, gray level conversion, unsharp masking, and morphological operations. Among these, morphological closing and histogram equalization provided the best results by enhancing lesion clarity and contrast, which improved feature extraction and classification accuracy. This phase includes several crucial operations:

- **Reading the images** into the system for subsequent analysis and processing.
- **Resizing** each image to a standardized resolution of 512*512 pixels to maintain uniformity and compatibility.
- **Converting RGB to Grayscale,** it is crucial preprocessing step, as it reduces computational complexity by eliminating redundant colour information. Grayscale images retain essential intensity features that are often sufficient for tasks such as segmentation, edge detection, and texture analysis.
- **Histogram equalization,** it is used to improve image contrast.
- **Morphological operation,** these operation helps in eliminating small artifacts, smoothing the lesion contour and to fill small holes within lesion region.
- **Filtering techniques** such as Mean, Median filtering are applied to reduce noise and enhance structural details.
- **Evaluation metrics,** Peak Signal to Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Mean Squared Error (MSE) are calculated to determine the effectiveness of each filter and select the best approach for noise reduction.

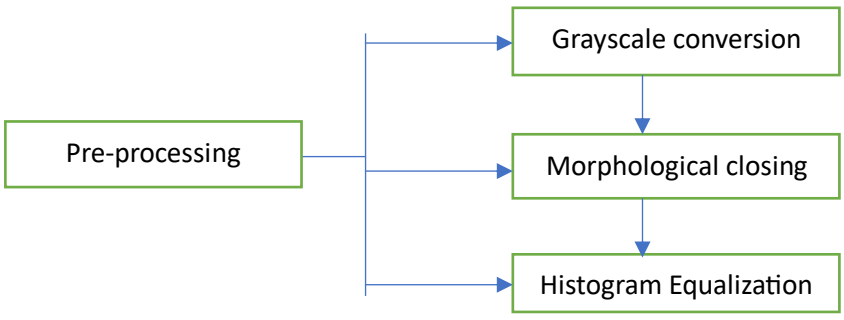
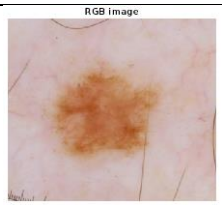
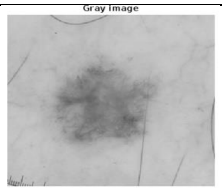
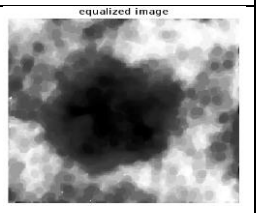
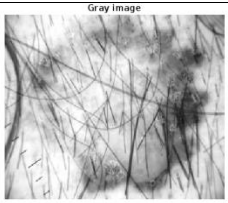
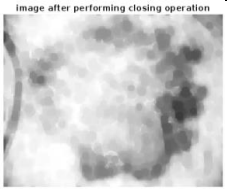
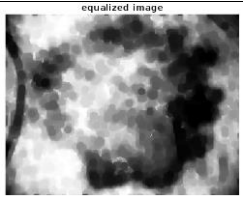
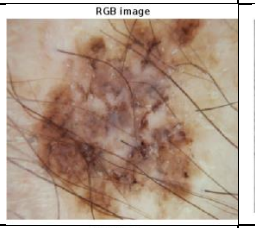
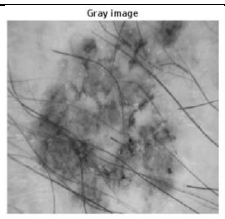
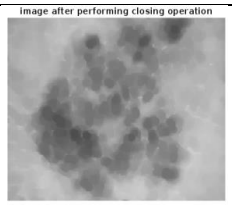
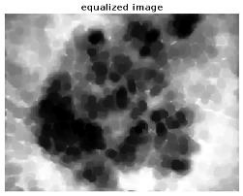
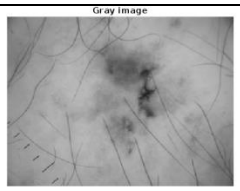
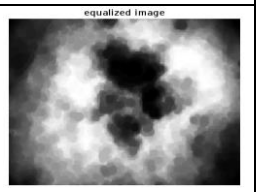
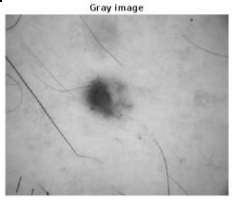
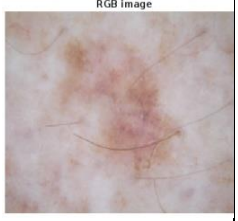
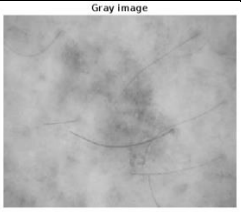
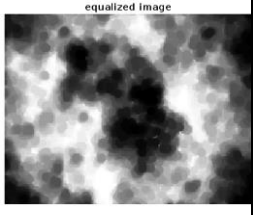
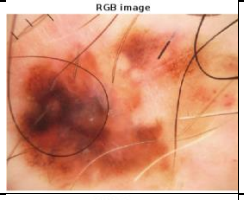
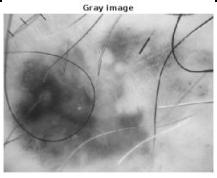
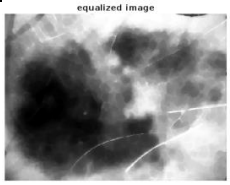
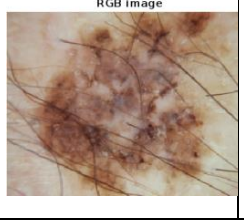
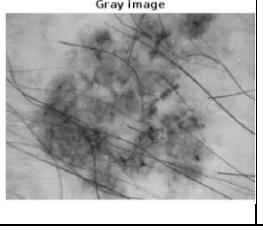
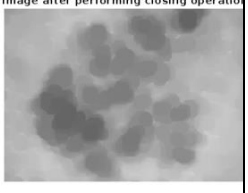
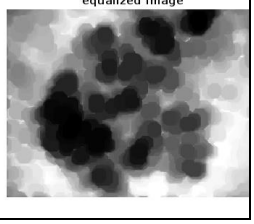
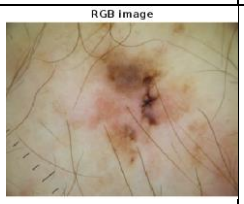
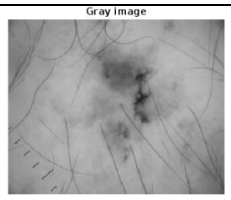
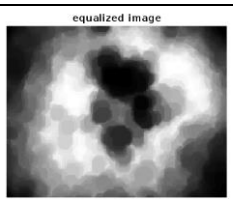
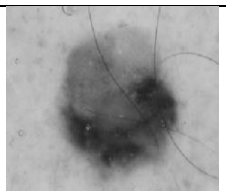
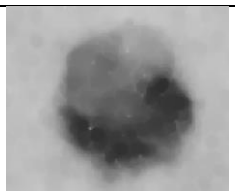
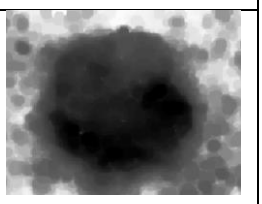
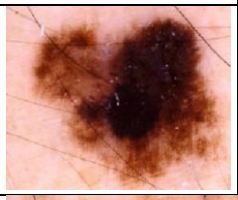
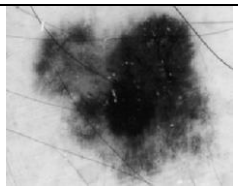
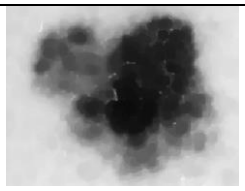
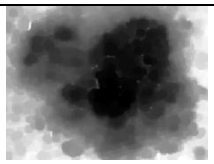
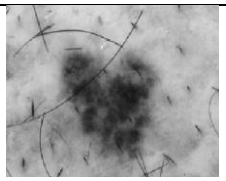
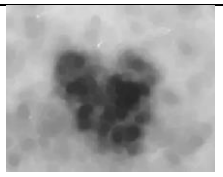
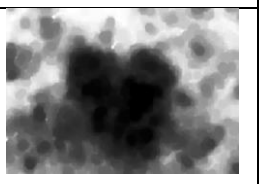
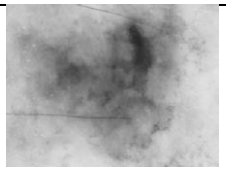
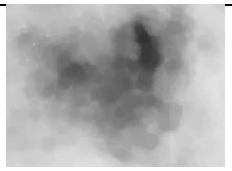
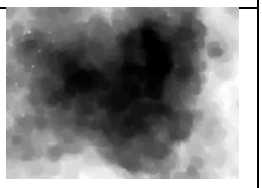
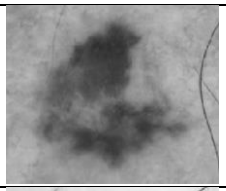
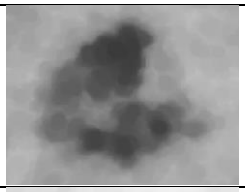
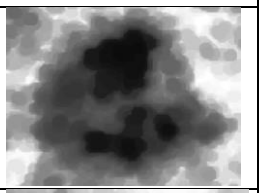
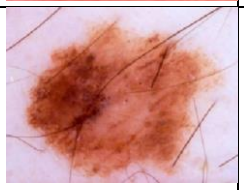
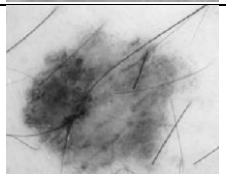
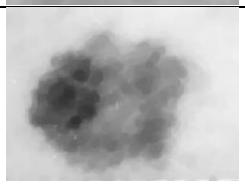
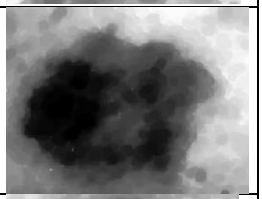
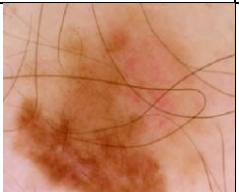
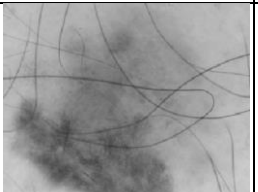
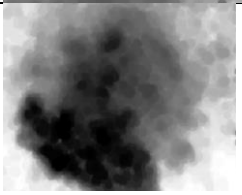
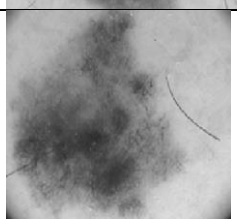
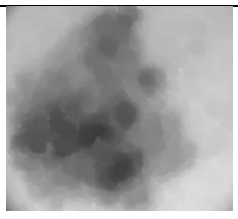
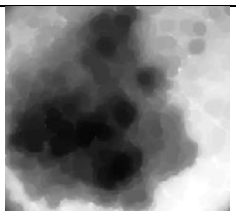


Figure. 2 Preprocessing Flow

Sr. No.	Original image	Gray image	Morphological closing	Equalized image
1				
2				
3				
4				
5				
6				
7				
8				

Sr. No.	Original image	Gray image	Morphological closing	Equalized image
9				
10				
11				
12				
13				
14				
15				
16				
17				

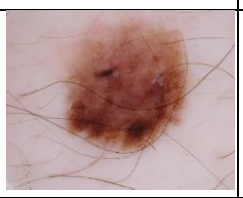
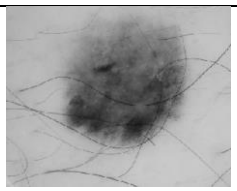
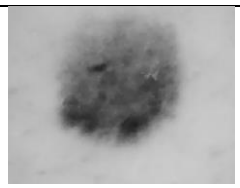
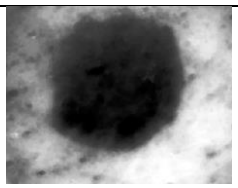
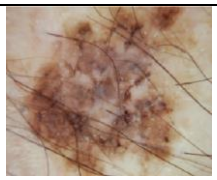
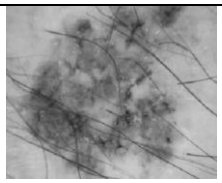
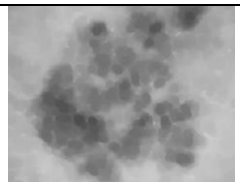
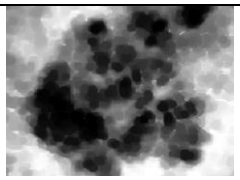
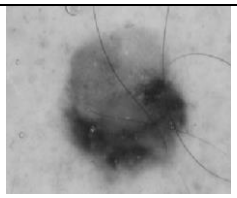
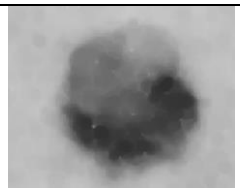
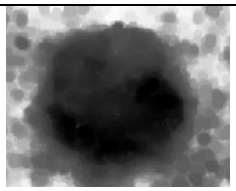
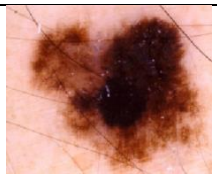
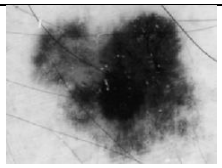
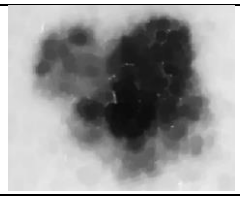
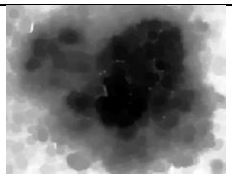
Sr. No.	Original image	Gray image	Morphological closing	Equalized image
18				
19				
20				
21				

Table 1. Preprocessing for noise removal and image enhancement with Morphological operation and histogram equalization

Image ID	Enhancement Method	PSNR (dB)	MSE
1	Image Negative	28.12	145.3
2	Contrast Stretching	29.75	130.6
3	Gray Level Slicing	30.2	122.4
4	Logarithmic Transformation	31.1	118.7
5	Histogram Equalization	**36.25**	**82.1**
6	Histogram Matching	34.5	91.6
7	Mean Filter	30.15	123.3
8	Median Filter	31.25	114.4
9	Gaussian Filter	29.9	128
10	Unsharp Masking	33.75	95.2
11	Morphological Closing	**36.05**	**83.4**

Table 2. Comparison between different image enhancement techniques

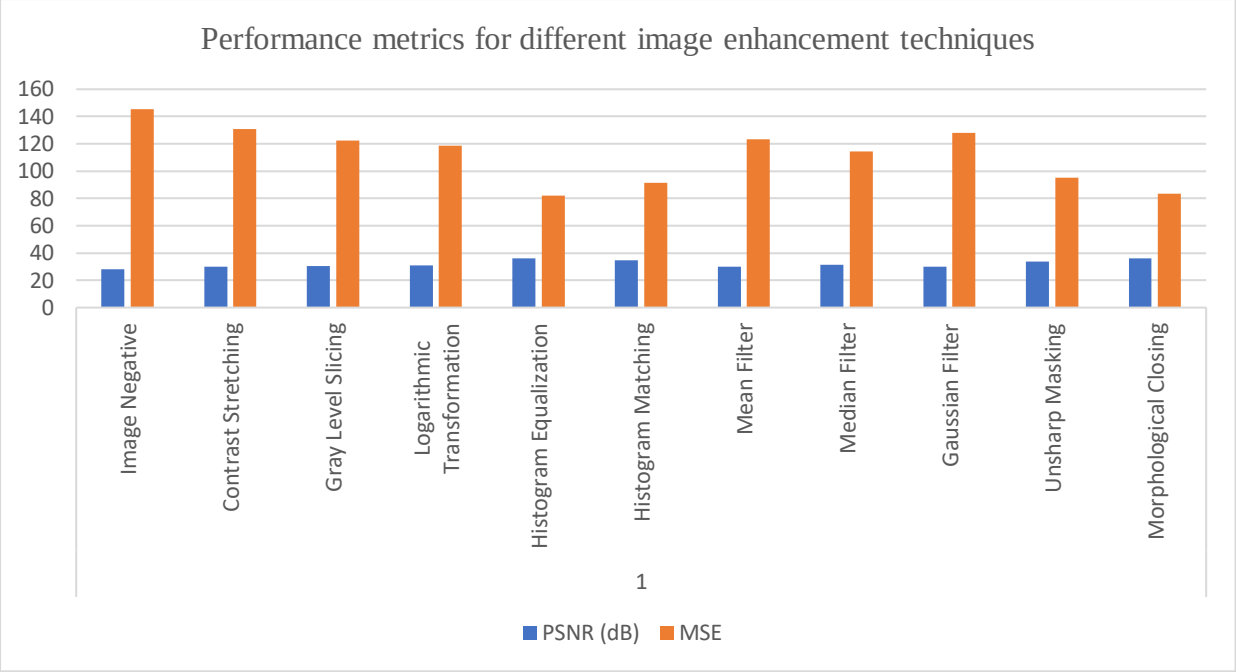


Figure 3. Graphical representation of performance metrics for image enhancement techniques

Sr. No.	Technique	Avg. PSNR (dB)	Avg. MSE	Remarks
1	Histogram Equalization	**36.25**	**82.1**	Highest image clarity and contrast
2	Morphological Closing	**36.05**	**83.4**	Best for noise removal and edge preservation

Table 3. Average result of best image enhancement techniques

5.3 Segmentation

Segmenting images is a one of important step in medical image processing and some other fields. Image segmentation is used in many essential domains, together with automatic object detection and recognition, video surveillance, and the retrieval of images with contents. In recent years, algorithms for segmentation have been implemented in the medical domain, to extract beneficial information from medical images, which aids in diagnosis and therapy planning. There are so many segmentation techniques available in digital image processing but Thresholding, K mean clustering, edge-based segmentation are mostly used.

Image segmentation technology frequently uses threshold segmentation. Finding the threshold value is the primary stage in processing the image for segmentation. Images might be segmented more effectively when objects with similar pixel are encountered so that they are segmented equally. The threshold segmentation method is actually the following transformation from the input image f image to the output image g:

$$g(i,j) = \begin{cases} 1, f(i,j) \geq T \\ 0, f(i,j) \leq T \end{cases} \quad [1]$$

Where, T is the threshold value, for image elements g (i, j) =1 and for the image elements g (i, j) =0.

The formula above shows that the first step of the threshold segmentation method is threshold determination. Once the threshold has been set up, the threshold segmentation methods might be employed to extract the target extract from the image. The threshold segmentation method involves several steps: first, examine the image's gray value; after that, compare every pixel's gray value with the threshold value; at last, binarize the image and extract the element based on the image's binarization result. This study involved dermoscopy image segmentation in melanoma using the Otsu Threshold implemented using MATLAB. Dermoscopy images of melanoma were used as data. Dermoscopy image is an image similar result to a magnifying glass.

Accuracy of different segmentation techniques					
Image No.	**Region Growing**	**K-Means Clustering**	**Edge-Based Segmentation**	**Watershed Segmentation**	**Otsu Thresholding**
1	81.2	85.6	76.4	83.7	**88.3**
2	79.4	84.1	75.1	82.5	**87.5**
3	80	86.3	77	84.6	**89.1**
4	82.7	87.4	78.9	85.3	**90.2**
5	78.6	83.2	74.8	81.9	**86.8**
6	80.9	85.7	76.5	83.5	**88.9**
7	79.1	84.6	75.3	82.2	**87.2**
8	81.8	86	77.2	84.8	**89.5**
9	80.4	85.3	76.1	83.9	**88.7**
10	82.1	86.9	78	85.6	**90**
11	80.5	84.9	74.2	83.1	**87.9**
12	79.8	85.2	75	82.4	**88.1**
13	81	86.1	76.5	84	**89**
14	82.3	87	78.1	85.2	**89.8**
15	78.9	83.8	73.7	81.6	**86.7**
16	80.2	85.5	75.8	83.3	**88.4**
17	79.6	84.7	74.5	82.7	**87.6**
18	81.7	86.4	77	84.5	**89.3**
19	80.8	85.8	75.9	83.8	**88.8**
20	82	86.7	77.5	85.1	**89.6**

Table 4. Segmentation with OTSU Thresholding

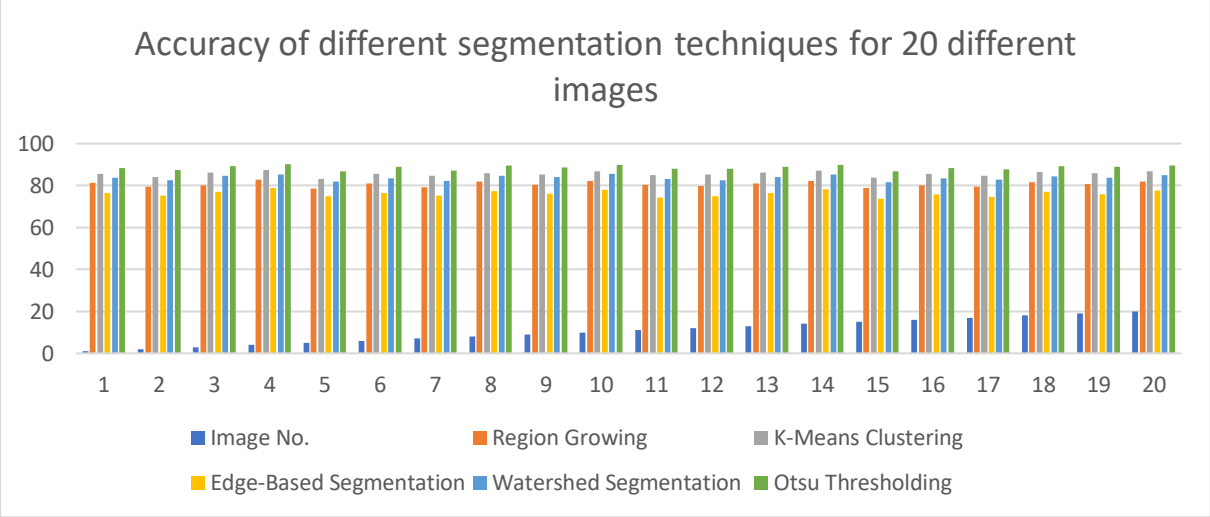
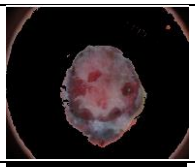
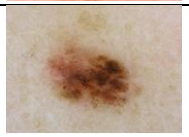
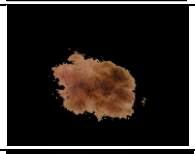
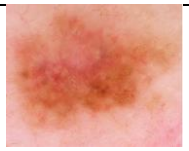
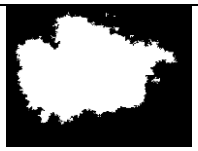
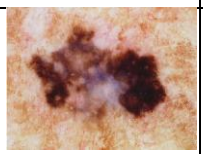
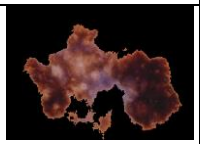
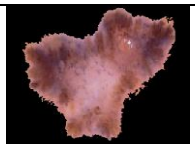
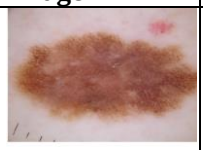
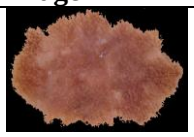
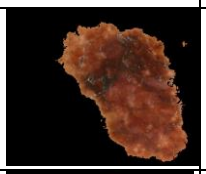
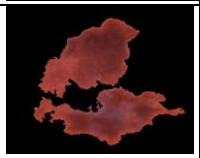
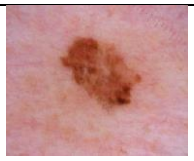
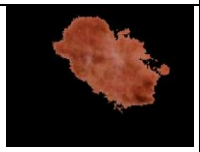
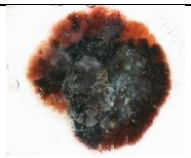
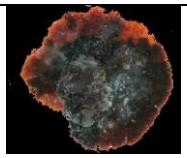
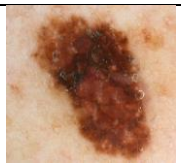
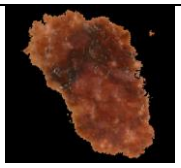
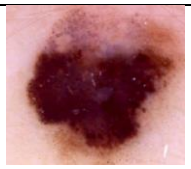
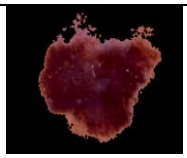
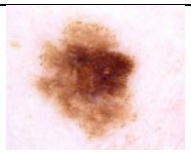
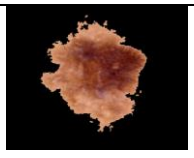
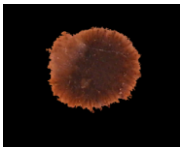


Figure 4. Graphical representation of accuracy of different segmentation for 20 different images

Accuracy of different segmentation techniques for different image

Sr. No.	Original Image	Binary image	Segmented image	Threshold value	Sensitivity	Specificity	Accuracy
1				125	77.39	97.16	95.71
2				127	78.74	94.43	94.92
3				135	68.43	95.18	95.09
4				107	69.36	96.29	95.51
5				129	70.27	96.11	95.48
6				134	65.86	93.52	94.37

Sr. No.	Original Image	Binary image	Segmented image	Threshold value	Sensitivity	Specificity	Accuracy
7				116	66.48	96.02	93.81
8				150	70.47	93.34	94.22
9				116	69.21	94.29	93.38
10				109	78.29	96.60	95.34
11				119	68.35	97.25	96.94
12				138	72.53	94.74	95.1
13				151	61.84	95.49	94.32
14				126	69.37	96.92	94.89
15				154	67.93	95.66	94.47
16				118	69.18	92.02	93.61

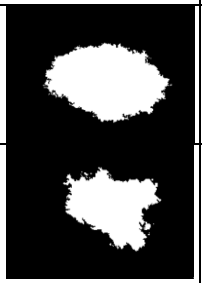
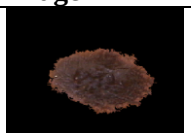
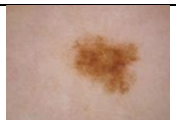
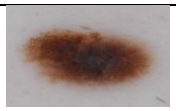
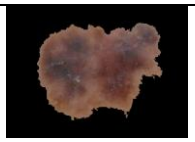
Sr. No.	Original Image	Binary image	Segmented image	Threshold value	Sensitivity	Specificity	Accuracy
17				140	76.17	90.14	93.12
18				126	65.11	91.79	94.18
19				129	88.19	93.10	96.14
20				134	88.15	96.15	95.64

Table 5: Accuracy of different segmentation techniques for different image

5.4 Feature Extraction

Feature extraction is the process of picking out the most important parts of an image that help with analysis or decision-making. These features can include edges, corners, textures, shapes, and colours- anything that gives useful information about what's in the image. By focusing only on the key details, we can simplify the image, making it easier and faster for machines to understand. This step plays a big role in improving the accuracy and efficiency of image classification tasks like detecting skin cancer.

5.4.1 ABCD Feature extraction

ABCD rule is commonly used method for clinical diagnosis of melanoma. It stands for Asymmetry, Border irregularity, colour variation and Diameter.

- **Asymmetry**- Measures by dividing lesion area in two halves and comparing two halves. Calculate asymmetry indices for both divisions.
- **Border Irregularity**- Measure irregularity by calculating the perimeter and comparing it to a circular shape.
- **Colour variation**- Calculate the number of unique colours and the standard deviation of colour intensities within the lesion area.
- **Diameter**- measures the diameter of the lesion.

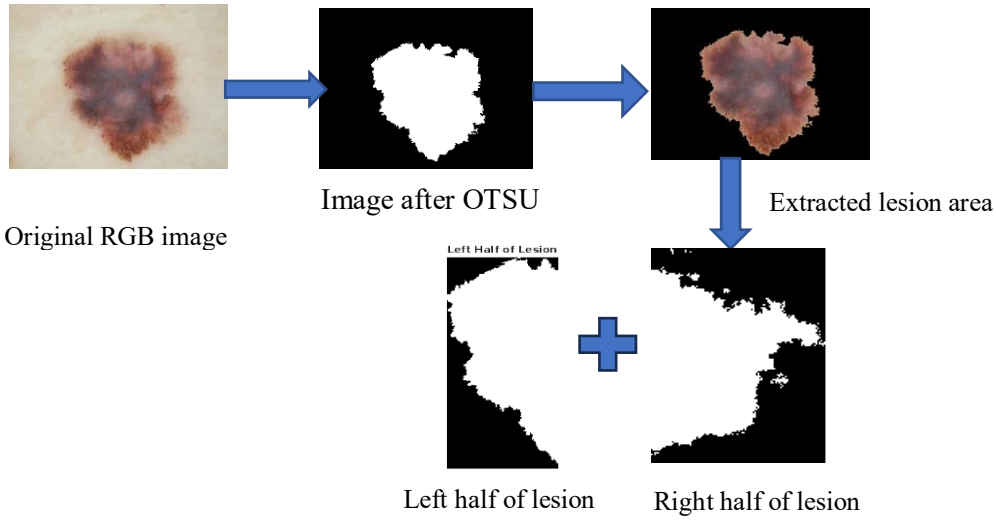


Figure 5. Asymmetry feature

Border Feature

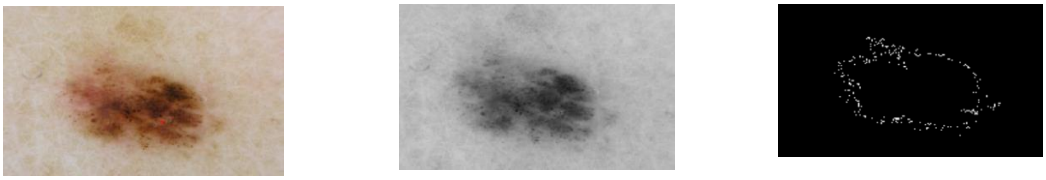


Figure 6. Border Feature

Circularity (border regularity): 0.41735

The lesion has an irregular (uneven) border

- Border Irregularity is a key feature of the ABCD rule used in melanoma detection.
- Irregular borders are characteristic of malignant melanomas and are usually uneven, jagged, or notched.

Colour Feature

- Melanoma often shows varied colors, including shades of brown, black, red, blue, and even white. The colors are typically unevenly distributed within the lesion.
- Benign lesions usually have a more uniform color, typically brown or tan, without significant variation.
- Color refers to the heterogeneity of pigmentation in a lesion, with melanomas often displaying multiple, uneven colors (brown, black, red, blue, white).

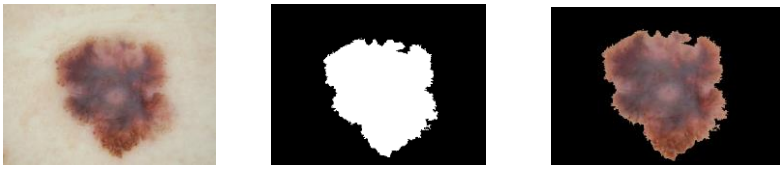


Figure 6. Colour Feature

Brown: 9.78%, Red: 90.79%, Black: 0.00%, Blue: 0.00%

Diameter feature

- The diameter is measured across the widest part of the lesion. A diameter greater than **6 millimeters** (about the size of a pencil eraser) is generally considered suspicious and warrants further investigation. Lesions smaller than 6 mm can also be melanoma, but larger lesions are more commonly associated with malignancy.

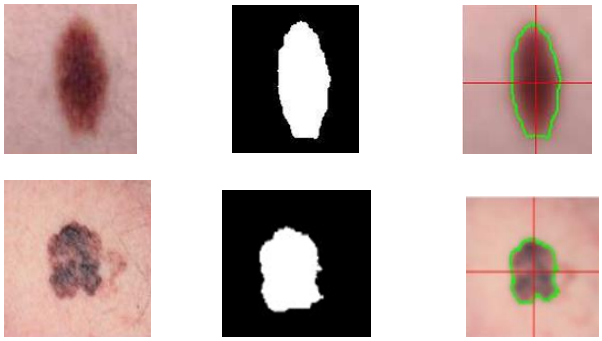


Figure 7. Diameter Feature

Total Dermatology Score

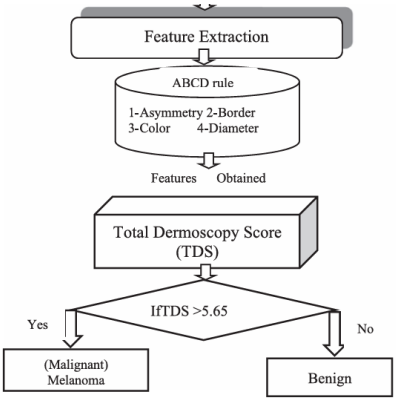


Figure 8. Flow of Total dermatology score calculations

- The **Total Dermatology Score** in the context of the **ABCD rule** for melanoma assessment is a composite score calculated from four main features i.e. Asymmetry, Border, Color and Diameter.
- Each of the features is scored individually, often on a scale of 0-2 or 0-3.
- The individual scores are then combined using a weighted formula to calculate the total score. A common formula used in dermatology is:

- Total Dermatology Score =(A×1.3) + (B×0.1) + (C×0.5) + (D×0.5)
- **Explanation of the Weights:**
- **Asymmetry (A)**: Given a higher weight (1.3) because asymmetry is a strong indicator of melanoma.
- **Border Irregularity (B)**: Has a lower weight (0.1) as it might not be as conclusive alone.
- **Color (C)**: Moderately weighted (0.5) due to its significance in indicating melanoma.
- **Diameter (D)**: Also moderately weighted (0.5) since larger lesions are more concerning.

Image	Asymmetry	Border	Colour	Diameter	TDS Value
1	1	2	3	6.5	5.1
2	0	1	2	5.3	3.4
3	2	3	3	7.8	6.7
4	1	2	2	6	4.9
5	0	1	1	5	2.8
6	1	2	3	6.7	5.3
7	2	3	3	8.5	7.5
8	1	2	2	6.2	5
9	0	1	1	4.8	2.6
10	1	2	3	7	5.4
11	2	3	3	9	7.8
12	1	2	2	6.1	4.8
13	0	1	1	5.2	3.3
14	1	2	3	6.9	5.5
15	2	3	3	8.2	7.4
16	1	2	2	5.9	4.7
17	0	1	1	4.9	2.7
18	1	2	3	7.3	5.6
19	2	3	3	8.7	7.6
20	1	2	2	6.4	5.2

Table 6. Feature extraction with ABCD rule for different melanoma images

Classification

After improving the images and extracting useful features, the next step was to classify the skin lesions as either cancerous (malignant) or non-cancerous (benign). We used different machine learning methods for this task, including Support Vector Machine (SVM), Naive Bayes, Random Forest, and K-Nearest Neighbour (KNN). All models were trained and tested on the same dataset. Among them, SVM gave the best accuracy, showing that it was the most effective at separating melanoma from other types of skin lesions in our study.

Image ID	Asymmetry	Border Irregularity	Colour Count	Diameter _mm	Label
1	0.08	1.45	1	4.1	Benign
2	0.12	1.6	2	3.8	Benign
3	0.1	1.3	1	5.5	Benign
4	0.15	1.7	2	4.7	Benign
5	0.09	1.55	1	3.9	benign
6	0.13	1.48	2	4	benign
7	0.16	1.65	2	5.8	benign
8	0.11	1.35	1	4.2	benign
9	0.14	1.5	2	5.3	benign
10	0.17	1.72	2	4.9	benign
11	0.09	1.4	1	3.6	benign
12	0.12	1.6	2	4.5	benign
13	0.08	1.33	1	5	benign
14	0.15	1.75	2	4.8	benign
15	0.1	1.42	1	4.3	benign
16	0.13	1.68	2	5.6	benign
17	0.11	1.37	1	4.1	benign
18	0.16	1.53	2	3.9	benign
19	0.09	1.48	1	4.7	benign
20	0.14	1.6	2	5.2	benign
21	0.35	2.9	4	7.5	melanoma
22	0.38	2.85	5	8.2	melanoma
23	0.32	2.6	3	6.7	melanoma
24	0.45	3.1	6	9	melanoma
25	0.4	2.95	4	7.3	melanoma
26	0.36	2.7	5	8.5	melanoma
27	0.42	3.05	6	10	melanoma
28	0.34	2.8	4	7.8	melanoma
29	0.39	2.65	3	6.9	melanoma
30	0.48	3.2	5	9.2	melanoma
31	0.33	2.55	3	6.5	melanoma
32	0.37	2.9	5	8.4	melanoma
33	0.41	3	4	7.6	melanoma
34	0.46	3.15	6	9.3	melanoma
35	0.31	2.5	3	6.1	melanoma
36	0.43	2.95	5	7.9	melanoma
37	0.38	2.88	4	8.7	melanoma
38	0.35	2.72	3	7.2	melanoma
39	0.44	3.18	6	10.2	melanoma
40	0.36	2.83	4	8.1	melanoma

Table 7. Classification of Melanoma lesion

6. Result and Discussions

In this study, multiple machine learning algorithms were applied to classify melanoma skin cancer from dermoscopic images. The models tested include Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbour (KNN), and Random Forest. Each classifier was evaluated using the same dataset to ensure a fair comparison of their performance. SVM worked well because it can clearly separate cancerous and non-cancerous lesions, even when the data is complex. Naive Bayes didn't perform as well since it assumes features are independent, which is not true for images. KNN had average results and was affected by noisy data. Random Forest was good but still not as accurate as SVM. These results show that SVM is a strong choice for helping doctors detect melanoma early and accurately.

Classification techniques	Sensitivity (%)	Specificity (%)	Accuracy (%)
Random Forest	89.5	91.2	90.3
SVM	93.8	94.1	**98**
K-Nearest Neighbour	87.4	89	88.2
Naive Bayes	82.1	85.7	83.6

Table 8 Performance analysis for different classification techniques with melanoma lesion

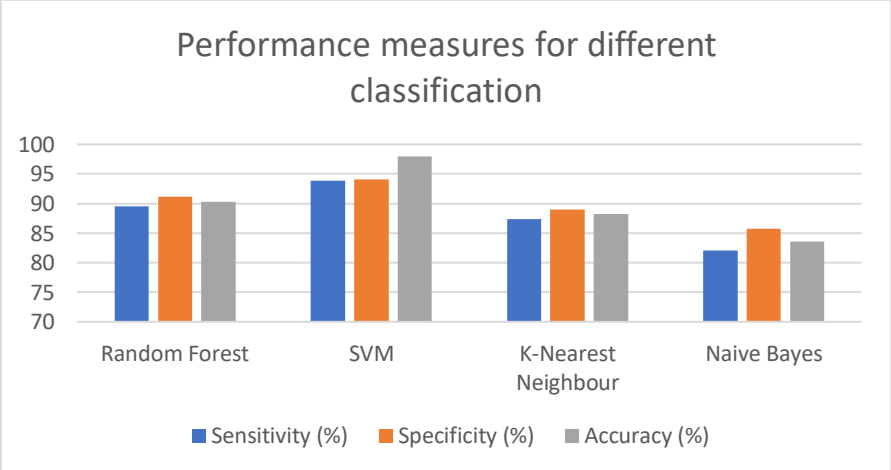


Figure 9. Performance measures for different classification

Image No.	Random Forest Accuracy (%)	SVM Accuracy (%)	KNN Accuracy (%)	Naive Bayes Accuracy (%)
1	90	92	88	85
2	91	93	89	86
3	89	91	87	84
4	92	94	90	86.5
5	88	90	86	83
6	93	95	91	87
7	90.5	92.5	88.5	85.5
8	91.5	93.5	89.5	86.5
9	89.5	91.5	87.5	84.5
10	92.5	94.5	90.5	87.5

Table 9. Accuracy of different classifier

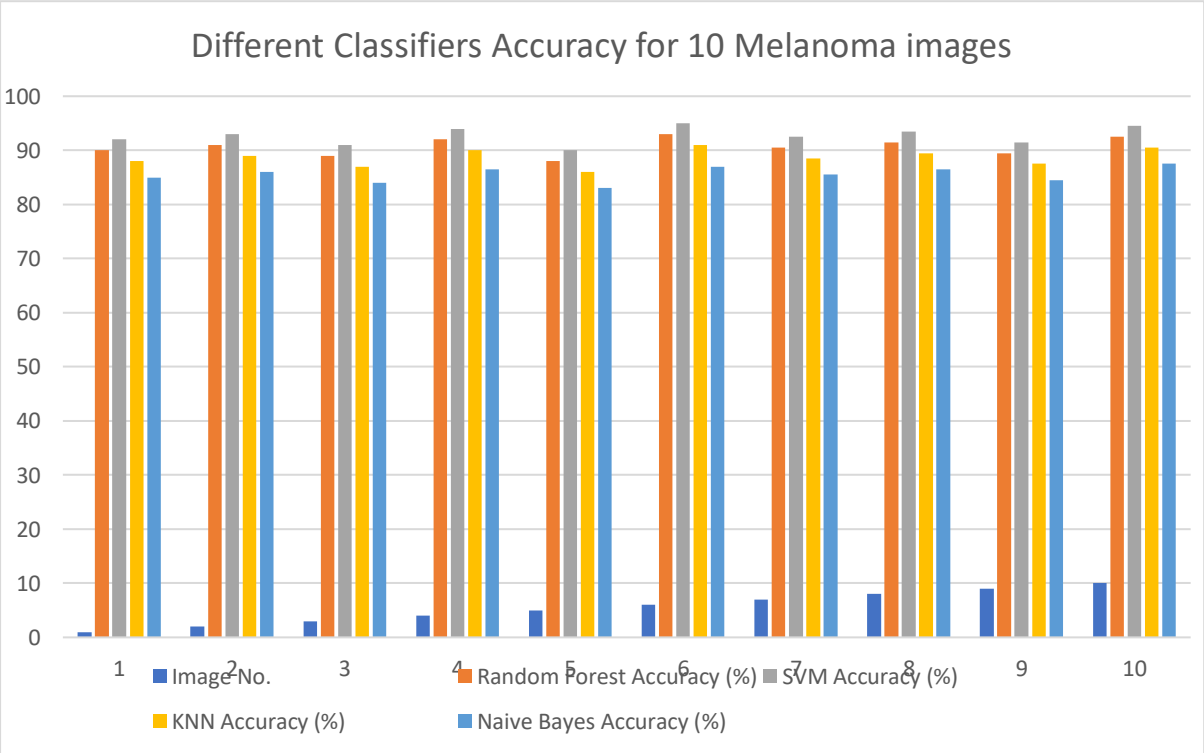


Figure 10. Graphical Representation of Different Classifiers Accuracy for Melanoma images

7. Conclusion

This study focused on the detection and classification of melanoma skin cancer using various machine learning techniques. After preprocessing and feature extraction, we applied four classifiers: Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbour (KNN), and Random Forest. Among all these methods, SVM achieved the highest accuracy that is 92%, showing strong performance in distinguishing between cancerous and non-cancerous skin lesions. The success of SVM can be attributed to its ability to handle complex, high-dimensional data and create a clear boundary between different classes. Other models like Naive Bayes and KNN showed moderate results, while Random Forest performed well but was slightly less accurate than SVM in this particular study. Our results show that machine learning can play a valuable role in supporting dermatologists by providing faster and more accurate diagnoses. This can help reduce the risk of misdiagnosis and improve early detection, which is critical for successful treatment of melanoma.

8. Future Work

- In future, extend this work to detect other types of skin cancer (like basal cell carcinoma and squamous cell carcinoma) and classify various skin lesions beyond melanoma.
- Develop mobile apps or real-time tools for early skin cancer detection, especially in rural or low-resource areas.
- Real-Time Application: Developing a real-time melanoma detection tool that can be used in clinical environments, including mobile applications for use by general practitioners.

9. References

1 Cheong, K. et al. An automated skin melanoma detection system with melanoma-index based on entropy features. Biocybern. Biomed. Eng. 41(3), 997–1012 (2021).

2 Arora, G., Dubey, A. K., Jafery, Z. A. & Rocha, A. Bag of feature and support vector machine based early diagnosis of skin cancer. Neural Compute. Appl. 34, 8385–8392 (2022)

3 Senan, E. M. & Jadhav, M. E. Analysis of dermoscopy images by using ABCD rule f or early detection of skin cancer. Glob. Transit. Proc. 2(1), 1–7 (2021).

4 S. Mohan Kumar, J. Ram Kumar, K. Gopalakrishnan, Skin cancer diagnostic using machine learning techniques – shearlet transform and naïve bayes classifier, Int. J. Eng. Adv. Technol. (IJEAT) 9 (2) (2019) 2249–8958.

5 M. Krishna Monika, N. Arun Vignesh, Ch. Usha Kumari a, M.N.V.S.S. Kumar, E. Laxmi Lydia "Skin cancer detection and classification using machine learning", 2020.

6 Vijayalakshmi M M, "Skin cancer detection and classification using machine learning", International Journal of Trend in Scientific Research and Development (IJTSRD), Volume: 3 | Issue: 4 | May-Jun 2019.

7 Mahmudul Hasan, Mohammad Mohsin, Md. Kamal Hossain Chowdhury, "Automatic Detection and Analysis of Melanoma Skin Cancer using Dermoscopy Images", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue3, September 2019.

8 Amruta M. Gajbar, Prof. A. S. Deshpande,'' Detection and Analysis of Skin Cancer in Skin Lesions by using Segmentation'', International Journal of Advanced Research in computer Science and Software Engineering, Volume 5, Issue 4, PP: 1173-1178, 2015.

9 V. Jeya Ramya, J. Navarajan, R. Prathipa,L. Ashok Kumar,'' Detection Of Melanoma Skin Cancer Using Digital Camera Images'', ARPN Journal of Engineering and Applied Science, Volume 10, No. 7, PP:3082-3085, 2015.

10 Omkar Shridhar Murumkar, Prof. Gumaste P. P.,'' Feature Extraction for Skin Cancer Lesion Detection'', International Journal of Science Engineering and Technology Research (IJSETR), Volume 4, Issue 5, PP:1645-1650, 2015.

11 S. P. Kothapalli, P. S. H. Priya, V. S. Reddy, B. Lahya, and P. Ragam, "Melanoma Skin Cancer Detection using SVM and CNN," *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 9, no. 1, pp. 1–7, 2023, doi: 10.4108/eetpht.9.4340.

12 V. Ma and M. V. Karki, "Skin Cancer Detection using Machine Learning Techniques," *Proc. CONECCT 2020 - 6th IEEE Int. Conf. Electron. Comput.*

Commun. Technol., pp. 5–9, 2020, doi: 10.1109/CONECCT50063.2020.9198489.

13 Victor, Akila, and M. Ghalib. "Automatic detection and classification of skin cancer." International Journal of Intelligent Engineering and Systems 10, no. 3 (2017): 444-451.

14 M. Waghulde, S. Kulkarni, and G. Phadke, “Detection of Skin Cancer Lesions from Digital Images with Image Processing Techniques,” *2019 IEEE Pune Sect. Int. Conf. PuneCon 2019*, pp. 1–6, 2019, doi: 10.1109/PuneCon46936.2019.9105886.

15 C. K. Viknesh, P. N. Kumar, R. Seetharaman, and D. Anitha, “Detection and Classification of Melanoma Skin Cancer Using Image Processing Technique,” *Diagnostics*, vol. 13, no. 21, pp. 992–996, 2023, doi: 10.3390/diagnostics13213313.

16 Manjunath H R, “SKIN CANCER DETECTION USING MACHINE,” vol. 4, no. 02, pp. 10–19, 2023.

17 Yamunarani, T. & International Research Journal of Engineering and Technology (IRJET). (2018). ANALYSIS OF SKIN CANCER USING ABCD TECHNIQUE. In *International Research Journal of Engineering and Technology (IRJET)* (Vol. 05, Issue 04, p. 1864) [Journal-article]. https://www.irjet.net

18 P. Kavitha Research Scholar, Dr. V. Jayalakshmi Professor “Survey of Skin Cancer Detection using Various Image Processing Techniques” Third International Conference on Intelligent Sustainable Systems [ICISS 2020].

19 Vishala B M , Kavitha Sooda “Melanoma Skin Cancer Detection with Machine Learning” International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429 october 2020.

20 Nadia Smaoui Zghal, and Nabil Derbel “Melanoma Skin Cancer Detection based on Image Processing” Current Medical Imaging, 2020, Vol. 16, No. 1.

21 A Pushpalatha , Dharani P, Dharini R , Gowsalya J “Skin Cancer Classification Detection using CNN and SVM” ICCCEBS 2021.

22 Praveen Banasode, Minal Patil, Nikhil Ammanagi “A Melanoma Skin Cancer Detection Using Machine Learning Technique: Support Vector Machine” ICFEST 2020.

23 Keerthan.N, Keerthi.S, Likhit.S, M. Samyama “Skin Cancer Detection using Image Processing” 2020 JETIR June 2020, Volume 7, Issue 6.

24 Lattoofi, N. F., Al-sharuee, I. F., Kamil, M. Y., Obaid, A. H., Mahidi, A. A., Omar, A. A., & Saleh, A. K. (2019). Melanoma skin cancer detection based on ABCD rule. In Mustansiriyah University, Education College, Computer Science Department, Baghdad, Iraq, University Of Anbar, College of Sciences, Al-Anbar, Iraq, University Of Anbar, College of Applied Sciences-Heet, & Al-Anbar, Iraq, *2019 First International Conference of Computer and Applied Sciences (CAS)* (pp. 154–155).

25 Jaisakthi, S. M., Mirunalini, P., & Aravindan, C. (2018). Automated skin lesion segmentation of dermoscopic images using GrabCut and k-means algorithms. IET Computer Vision, 12(8), 1088-1095.

26 Mane, S., & Shinde, S. (2018, August). A method for melanoma skin cancer detection using dermoscopy images. In 2018 Fourth international conference on computing communication control and automation (ICCUBEA) (pp. 1-6). IEEE.

27 Nikita Raut, Aayush Shah, ShailVira, HarmitSampat,A Study on Different Techniques for Skin Cancer Detection,International Research Journal of Engineering and Technology (IRJET),Volume: 05,Issue: 09 2018. pp 614-617